Description of Generic Sensor Types for the Continuous Diagnostic and Mitigation (CDM) Collection System

May 2014

1 Purpose

The Continuous Diagnostics and Mitigation (CDM) Program uses sensors[1] to collect desired state and actual state data to identify defects. Each CDM capability (e.g., Hardware Asset Management) will incorporate sensors specific to the collection of data necessary to identify defects for that capability. While the same sensor will often support multiple capabilities, what it collects and provides may be different for each one. More information about how sensors are incorporated into the architecture for the Collection System is documented in *Continuous Diagnostic and Monitoring (CDM) Capability Module Generic Architecture.* This document is intended to describe the generic sensor types listed in the generic architecture, to include information about potential for operational impacts and data accuracy associated with a particular sensor type. A summary table is included for the actual state sensor types.

Capability-specific documents will provide more detailed information about generic sensor types with respect to that capability. They will discuss actual state data accuracy issues specific that capability as well as provide an illustrative example of how data from different sensor types could be combined to address those issues.

[1] The term sensor is used very broadly to define anything (including a person) that senses, queries, contains, or provides data to CDM.

2 Generic Actual State Sensor Types

There are five generic sensor types defined for the collection of actual state data: active network sensor, passive network sensor, asset[2] management repository, network event sensor, and endpoint-based sensor. False positives are when a condition is identified in the operational environment but not actually present (e.g., a device is identified as having a particular software product installed when that product is not installed on the device). False negatives are when a condition is not identified but is present in the operational environment (e.g., a device is deployed and communicating on the network but not detected). The examples provided are not intended to be comprehensive; they are intended to provide insight into the types of entities that could be considered a sensor of that type.

2.1 Active Network Sensor

An Active Network Sensor actively probes the network or a device over the network. By utilizing credentials, Active Network Sensors are able to scan and collect accurate information from devices on the network. They are less prone to false positive reporting due to the fact the data being collected by the sensor is originating from the source device (i.e., data is not being proxied). False positives tend to occur when the sensor is not authorized to receive the information and defaults to a “fail” for a check or when the device being probed has been compromised and returns false information. Any false negatives experienced by Active Network Sensors are mostly the result of not being able to connect to a device because of network policies or controls (e,g, blocked by firewalls) or because the IP space is too large to be scanned in the allotted time period (e.g., IP v6). Non-credentialed sensors of this type will generate false positives if the sensor is not authorized to receive the information and defaults to a “pass” for a check. Active Network Sensors actively probe the network and devices to collect information, and this query and respond activity can impact certain constrained environments. Active probing can also have a negative impact on certain devices, causing them to fail or reboot. All of the above should be considered when deciding on how to deploy and use Active Network Sensors. Examples of Active Network Sensors include:

- Network scanners used to detect devices and collect device identification data (e.g. IP, hostname, SWID)
- Configuration Management Software Suites/Tools used to push updates and check for installed software packages and executable files/programs
- Vulnerability Management Suites/Scanners designed to assess devices and perform patch remediation, compliance auditing, and malware detection

2.2 Passive Network Sensor

A Passive Network Sensor is designed to capture network traffic that passes across a monitored network link. Passive Network Sensors are configured to collect and/or inspect communications that traverse a certain network segment to provide information about network connections and/or specific content of those connections. Due to their passive nature, these sensors are not always able to collect specific enough information (e.g., version or patch level of software installed on a device) from the typical communications it inspects. This will mainly result in reporting false positives, but under certain circumstance could result in false negatives as well. False negatives for Passive Network Sensors are mostly associated with passive sensors only being able to “see” data that traverses the network segment where it is located or configured to collect. While the collection of data by a passive sensor has no impact on the network, how it inspects, analyzes, and returns collection can potentially have latency and bandwidth implications. If inspection and/or collection are happening “in band” there can be latency issues. If strictly collection is occurring at the sensor and everything has to be sent back to a central server or collector for analysis then large amounts of data may have to be sent over the network at regular intervals. Examples of Passive Network Sensors include:

- Packet/Protocol analyzers (i.e. sniffers) used to detect device information transmitting on network segments
- NetFlow enabled devices (i.e. routers and switches) used to export records regarding IP traffic statistics over UDP

2.3 Asset Management Repository

An Asset Management Repository is a collection of data created and updated as part of a process or activity that manages that asset for an organization. It can also collect asset data acquired by proprietary or external means when the process of collection is not managed by the enterprise, but the results are made available (e.g., Mobile Device Management servers, clearance information for individuals employed by other agencies). Asset Management Repositories do not have to be incorporated into the CDM collection system, but can be used to provide timely and authoritative data directly to CDM. False positives (and possibly false negatives) generated by Asset Management Repositories are due to the repository maintaining information that is not in the proper format or at the necessary level of fidelity. This occurs because these repositories are part of an independent

[2] The term asset is used very broadly in this document to include any*thing* that the organization may need to manage from a security perspective (e.g., network, device, person)

activity that uses that data for its own purpose. False negatives will also occur because all assets that need to be checked may not be managed by this tool or process, causing data for some to not to be in the repository. Because these repositories are part of an existing collection system/process, the only bandwidth or accessibility issues are experienced on the "back end" (i.e., when information is provided to and/or accessed from CDM). If the data in the repository that is used by CDM is very dynamic, then there is more data to transmit more often. Sometimes there is more information in the repository than we need for CDM or the system and processes that manage the repository have unique security/confidentiality issues. All of these concerns need to be taken into account when deciding on how to get the information from the repository in an appropriate and timely manner to CDM. Examples of Asset Management Repositories include those that are used by:

- IP Management Software used to discover, monitor, and manage IP space for devices on network segments
- Mobile Device Management (MDM) services used to secure, monitor, and manage mobile devices
- Lightweight Directory Access Protocol (LDAP) used to serve as an information directory resource which allows the sharing of data about networks, users, services, and applications throughout the network
- Education and learning management system repositories that track employees and their training history
- Personnel security management systems that track the issuance of PIV cards
- Audit Management System that collect, aggregate, and analyze audit events from multiple devices

2.4 Network Event Sensor

A Network Event Sensor is designed to detect and report events of interest to a defined location in a timely manner. A Network Event Sensor differs from a Passive Network Sensor in that it is configured to not just collect information passing over the network segment being monitored, but to perform some level of analysis to identify predefined conditions and then act accordingly. The action might be to alert a person, aggregate events and send them to a server, or modify/block some part of the network communication. If detection and/or action are happening "in band" there can be latency issues that can negatively impact the network. These sensors are often configured to detect certain types of events, and then "tuned" to detect specific events of interest. In many cases the tuning of the sensor can be too narrow or too broad for a large amount of data. If the thresholds of a sensor are tuned too narrow events may be omitted from being triggered or alerted, while if the thresholds of a sensor are tuned too broad events can be missed during analysis or unable to be analyzed due to a large dataset. Both of these situations tend to result in false negatives. False positives can occur when signatures are developed that may be too broad for alerting purposes thus causing acceptable activities to be flagged as events by the Network Event Sensor. False positives also occur because the data available in the network communications may be inaccurate (i.e., the proxy's IP address is collected instead of the initiating device) or not specific enough. Better tuning of the sensor will drive down the false positives and false negatives of the sensor. Increased bandwidth on the network can occur if the thresholds and triggers are set to a level that identifies large numbers of events or requires that a large amount of data be passed back for events that are identified. Examples of a Network Event Sensor include:

- Security Information and Event Management (SIEM) Software/Tool used to gather, analyze, and present information about devices on the network
- Intrusion Detection System (IDS) used to monitor and detect events or policy violations on a network or device

2.5 Endpoint-Based Sensors

An Endpoint-Based Sensor is a software client installed on, or natively embedded within, the operating system of a device. Endpoint-Based Sensors have the lowest false positive rate due to the fact that the sensor is installed or embedded directly on the device. False positives can occur if reporting is not frequent enough or if the endpoint device becomes compromised and provides the sensor false data. False negatives for this type of sensor are directly related to whether or not the software client exists on the device and the periodicity and technique used to collect the information. If the information is published on a regular schedule or when a change occurs, then these types of false negatives are minimized. If an Active Network sensor needs to go collect the information, then this type of sensor will have the same false positive issues as those sensors. Bandwidth concerns are directly related to how much information is published or collected from the sensor and how often. Examples of Device Agents include:

- Security Endpoint Agents used to detect device anomalies, perform file integrity checks, and device system level analysis
- A Trusted Network Connect (TNC) client embedded in an endpoint OS

3 Generic Actual State Sensor Table

The following table summarizes the previous section. The columns of the table are defined as:

Type: The generic sensor type name.

Description: Basic description of the sensor type.

Weaknesses and Limitations: The general weaknesses or limitations of the sensor type with respect to the type of data it can collect.

False Positives: When data from this type of sensor may result in the reporting of an operational condition when that condition does not actually exist in the operational environment.

False Negatives: When data from this type of sensor may result in the non-reporting of an operational condition when that condition does exist in the operational environment.

Bandwidth Considerations: The typical impact on bandwidth that using this type of sensor can have on the operational environment.

Type	Description	Weaknesses and Limitations	False Positives	False Negatives	Bandwidth Considerations
Active Network Sensor	An Active Network Sensor actively probes the network or a device over the network.	Actively sensing large IP ranges takes extended amounts of time (possibly years)	False positives are rare for credentialed sensors because data is received by the sensor directly from the reporting endpoint device. Using non-credentialed sensors of this type result in higher false positive rates. False positives can also increase if the devices or software that responds to the sensor are compromised, spoofed, or modified to send false information (i.e.,	False negative rates are most directly associated with the sensor not being able to probe a set of devices. This can be the result of implementation issues (e.g., network connection policies) or overly large scan ranges (e.g., IPv6 blocks). Using non-credentialed sensors of this type that default to “pass” will also create false negatives.	Increased bandwidth overhead may occur from actively probing devices on the network as well as the high amount of responses received from devices.
Passive Network Sensor	Passive Network Sensors are configured to collect and/or inspect communications that traverse a certain network segment to provide information about network connections and/or specific content of those connections.	Limited network visibility due to subnetting and segmentation of the network Device state information can be inaccurate because the necessary level of fidelity is not provided as part of typical network communications	False positives are mostly related to the fact that the sensor is not able to determine specific enough information about a device (e.g., Software version and patch level) from data contained in typical network communications.	False negatives are almost exclusively due to sensor implementation issues since they can only see traffic that is traversing the same segment and only collect the data for which they are specifically configured.	Increased bandwidth can occur if large amounts of data must be passed back to a server for analysis.
Asset Management Repository	An Asset Management Repository is a repository created and updated as part of process or activity that manages that asset for an organization. It can also collect asset data acquired by proprietary or external means when the process of collection is not managed by the enterprise, but the results are made available.	Limited to only the devices identified for management by that tool or mechanism Information available is collected and used for a different purpose; so its limited to the data that it already provides in the format it already uses	False positives are due to the repository storing information in a format, at a level of fidelity, or with a context different from that which is needed.	False negatives are almost exclusively due to devices that are not managed as part of the larger capability, and therefore do not have data about them in the repository.	The only bandwidth increase is the “back end” updating or accessing of this stored information for use by CDM.

Type	Description	Weaknesses and Limitations	False Positives	False Negatives	Bandwidth Considerations
Network Event Sensors	A Network Event Sensor is designed to detect and report events of interest to a defined location in a timely manner. A Network Event Sensor differs from a Passive Network Sensor in that it is configured to not just collect information passing over the network segment being monitored, but to perform some level of analysis to identify predefined conditions and then act accordingly.	Associated device information can be inaccurate because the necessary level of fidelity is not provided as part of typical network communications The appropriate settings for thresholds and triggers are not easily defined or determined which can result in inaccurate identification of events	False positives mostly arise from improper tuning of the event thresholds or triggers. Definitions that are too broad often result in normal activities being flagged as events. The other main reason for false positives is that the data used to detect the event is not specific or accurate enough (e.g., actual device is behind a proxy).	False negatives are most commonly due to inadequate knowledge of triggers, inappropriate tuning of thresholds, and mishandling of events. Due to potentially large data sets, both thresholds and triggers that are too broad or too narrow will result in false negatives. If the scope is configured to be too broad with a large dataset, important events are missed or unable to be analyzed. If the scope is configured to be too narrow, the tuning prevents the identification of events.	Increased bandwidth can occur if the thresholds and triggers are set to a level that identifies large numbers of events or requires that a large amount of data be passed back for events that are identified.
Endpoint-Based Sensor	An Endpoint-Based Sensor is a software client installed on, or natively embedded within, the operating system of a device.	Potential negative impact on the device due to increased processor load	False positives only occur if reporting is not frequent enough or if the device is compromised and the agent has been modified to lie. The reporting frequency can be directly affected by configuration of the agent or indirectly affected by network connection issues.	False negatives are directly associated with what devices do not have the agent installed or software embedded.	Increased bandwidth overhead between device agent and collection manager. The amount of overhead is associated with how much information needs to be exchanged (e.g., heartbeat, only changes, full collection) and how often (e.g., every day, only when a change is detected).

Table 1: Generic Actual State Sensors Table

4 Generic Desired State Sensor Types

There are three generic sensor types for the collection of desired state data: Desired State Manager, Human Agent, and Management Repository

4.1 Desired State Manager

A Desired State Manager is a trusted individual who serves as the authoritative source for desired state sensor data. A Desired State Manager can directly enter, modify, and update desired state data in CDM. In the event that there is a conflict involving reported or collected data, the Desired State Manager must know where to get the authoritative answer for de-confliction and make the appropriate change.

4.2 Human Agent

A Human Agent is a trusted person who collects desired state information and enters it directly into CDM. Human Agents vary in expertise and experience, and therefore the quality of the data they collect and enter varies as well. Issues with data quality are usually related to the collection process. If the collection takes too long to complete or if it is prone to errors (i.e., due to the complexity of the process) then defects may be wrongly detected or missed based on inaccuracies in the desired state data.

4.3 Management Repository

A Management Repository is a collection of data created and updated by a process or activity that is responsible for managing that data. A Management Repository does not have to be incorporated into the CDM collection system to provide data directly to CDM. Because these repositories are maintained by tools and processes in the organization that are external to CDM, some information may be at the wrong specification level or in a different format than what is required for CDM. This can cause patterns of inaccuracies in the desired state data that can lead to systemic issues in defect identification. Examples of Management Repositories include those that are used by:

- Configuration Management systems that track all decision made by the organizations Configuration Control Board (CCB)
- Property Management systems that track device information like serial number, model number, and current location
- Human Resource systems that manage employee information like organization, current managers, and job title